CENTER FOR
MACHINE PERCEPTION

CZECH TECHNICAL
UNIVERSITY

REPRINT

Cheirality in Epipolar Geometry

Tomáš Werner and Tomáš Pajdla

{werner,pajdla}@cmp.felk.cvut.cz

T.Werner and T.Pajdla. Cheirality in Epipolar Geometry. *In Proc. of Eighth IEEE Int. Conference on Computer Vision, Vol. 1, pp. 548–553. Vancouver, Canada. July 7-14, 2001.*

Center for Machine Perception, Department of Cybernetics
Faculty of Electrical Engineering, Czech Technical University
Technická 2, 166 27 Prague 6, Czech Republic
fax +420 2 2435 7385, phone +420 2 2435 7637, www: http://cmp.felk.cvut.cz

Cheirality in Epipolar Geometry

Tomáš Werner Tomáš Pajdla
Center for Machine Perception
Faculty of Electrical Engineering, Czech Technical University
12135 Prague, Czech Republic
{werner,pajdla}@cmp.felk.cvut.cz

Abstract

The image points in two images satisfy epipolar constraint. However, not all sets of points satisfying epipolar constraint correspond to any real geometry because there can exist no cameras and scene points projecting to given image points such that all image points have positive depth. Using the cheirality theory due to Hartley and previous work on oriented projective geometry, we give necessary and sufficient conditions for an image point set to correspond to any real geometry. For images from conventional cameras, this condition is simple and given in terms of epipolar lines and epipoles. Surprisingly, this is not sufficient for central panoramic cameras. Apart from giving the insight to epipolar geometry, among the applications are reducing the search space and ruling out impossible matches in stereo, and ruling out impossible solutions for a fundamental matrix computed from seven points.

1. Introduction

It[1] is well-known that corresponding image points in two images of a 3-D scene satisfy the epipolar constraint, $\mathbf{x}^\top \mathbf{F}\mathbf{x}' = 0$ [3]. The less understood fact is that the reversed implication does not hold true – there exist configurations of image points satisfying the epipolar constraint that are not projections of any real 3-D scene. A simple reason is that if $\mathbf{x}$ is the first epipole and $\mathbf{x}'$ is not the second epipole, it is $\mathbf{x}^\top \mathbf{F}\mathbf{x}' = 0$ but no scene point projecting to $\{\mathbf{x}, \mathbf{x}'\}$ exists. A less trivial reason is that the reversed implication does not hold because the equation $\mathbf{x}^\top \mathbf{F}\mathbf{x}' = 0$ does not involve the positive depth constraint. In other words, the epipolar constraint alone allows some camera rays to intersect in front the cameras and some behind them.

[1]This research was supported by the Grant Agency of the Czech Republic under Project GACR 102/01/0971, EU Fifth Framework Programme Project OMNIVIEWS, No. 1999–29017, and by the Czech Ministry of Education under Research Programme MSM 212300013 Decision Making and Control in Manufacturing.

This fact has been addressed in several works. Hartley [4, 3] introduces the concepts of cheirality and quasi-affine transformation. The sign of cheirality (meaning in fact depth) $\chi(\mathbf{P}, \mathbf{X})$ indicates whether the scene point $\mathbf{X}$ lies in front of or behind the camera $\mathbf{P}$. For image correspondences to be projections of a real scene it is demanded that (i) the cheirality of all points be positive, (ii) the orientation of camera coordinate system bases be the same as for underlying true cameras. He shows that in this case the reconstructed scene differs from the true geometry by a quasi-affine transformation, which is a projectivity that preserves the convex hull of points. Werner *et al.* give an independent yet less exhaustive treatment of the topic [9].

Stolfi [7] develops a theory of oriented projective geometry. Oriented projective space T_n is defined as the set $\{\lambda\mathbf{x} | \mathbf{x} \in \mathbb{R}^{n+1}, \lambda > 0\}$ as opposed to projective space P_n which is the set $\{\lambda\mathbf{x} | \mathbf{x} \in \mathbb{R}^{n+1}, \lambda \neq 0\}$. Laveau and Faugeras [5] show the usefulness of oriented projected geometry for computer vision.

In multi-view geometry, usually only conventional TV cameras are considered. Recently, the advantages of *panoramic cameras* for various tasks have been recognized [1, 6]. These cameras have the viewing angle (at least at a single plane through its center) 360 degrees. Practical setups consisting of a curved mirror and a conventional camera have been proposed such that the whole device can be viewed as a camera with a single projection center. A panoramic camera with a single center is projectively equivalent to a conventional camera. However, if ray orientations are distinguished, a conventional camera brings an additional knowledge, namely the fact that all rays representing the image points are on a single side of a certain plane through it center, representing image line at infinity. Since this means that the camera looks in a certain direction, we call the conventional cameras *directional* [6].

Our contribution is as follows. We generalize the theory of cheirality and quasi-affine reconstruction for both directional and panoramic central linear cameras. Then the question is addressed whether two images are projections of any real scene by any real cameras. The most important result is the necessary and sufficient condition for two directional images, given in a simple form in terms of orientations of epipolar lines. Using Hartley's results [4], we prove that for directional cameras the sufficient condition is that either all epipolar lines aim towards the epipole or all aim out of it. Surprisingly, the directionality is crucial here – the condition is insufficient for panoramic images.

2. Used Concepts

We will consider images captured by central linear cameras. *Central* means that all camera rays intersect in a single point in space, the camera center. *Linear* means that the projection of a scene point to an image point is a linear mapping in homogeneous coordinates.

The central linear camera is usually sufficiently approximated by a conventional TV camera. A panoramic camera can be also made central and linear. If a panoramic camera consists of a conventional camera looking into a curved mirror with the shape of rotational hyperboloid or paraboloid, and if the mutual position of the mirror and the camera is appropriately adjusted, the device can be considered as a single central camera [1, 6].

An image captured by a central linear camera is a set of (oriented) directions in E_3 Euclidean 3-space. We model the *image space* by T_2. The elements of T_2 the image space are represented by homogeneous 3-vectors, $\mathbf{x} = (x_1, x_2, x_3)^\top \in \mathbb{R}^3, \mathbf{x} \neq \mathbf{0}$. Two vectors $\mathbf{x}$ and $\mathbf{x}'$ represent the same image point iff $\exists w > 0 : w\mathbf{x} = \mathbf{x}'$.

The field of view of a directional camera is contained in a hemisphere [6]. All vectors $\mathbf{x}_n$ representing the image points are on a single side of a single plane (image line at infinity), represented by $\mathbf{s}$. The field of view of a panoramic camera contains at least one great circle. For a panoramic image, there is no separating plane $\mathbf{s}$. If $\mathbf{s}$ is chosen as $\mathbf{s} = (0, 0, 1)^\top$, a directional image can be represented by a set $\{\mathbf{x}_n | \, \mathbf{x}_n \in \mathbb{R}^3, x_{n3} > 0, n = 1, \ldots, N\}$ of homogeneous 3-vectors with positive final coordinates. A panoramic image is represented by a set $\{\mathbf{x}_n | \, \mathbf{x}_n \in \mathbb{R}^3, n = 1, \ldots, N\}$ of arbitrary homogeneous 3-vectors. We define

Definition 1 *The image* $\{\mathbf{x}_n | \, \mathbf{x}_n \in \mathbb{R}^3, n = 1, \ldots, N\}$ *is directional if there exist* $\mathbf{s} \in \mathbb{R}^3$ *such that* $\forall n : (\mathbf{s}^\top \mathbf{x}_n > 0)$. *Otherwise, the image is panoramic.*

The *scene space* is in fact E_3. Since only its affine rather than metric properties are of interest in this paper, we will consider it as A_3 affine 3-space. A_3 is isomorphic to the part of T_3 lying in front of a certain plane, plane at infinity (horizon) Π_∞ [7]. This is denoted as $A_3 = (T_3, \Pi_\infty)$. We will represent Π_∞ by $\mathbf{\Pi}_\infty \in \mathbb{R}^4$ and choose $\mathbf{\Pi}_\infty = (0, 0, 0, 1)^\top$. Scene points are represented by homogeneous 4-vectors with positive final coordinate $\mathbf{X} = (X_1, X_2, X_3, X_4)^\top \in \mathbb{R}^4, \mathbf{X} \neq \mathbf{0}, X_4 > 0$. Sometimes, we will consider the scene space just as T_3 or P_3, dropping the constraint $X_4 > 0$. Two vectors $\mathbf{X}$ and $\mathbf{X}'$ represent the same scene point iff $\exists \xi > 0 : \xi\mathbf{X} = \mathbf{X}'$.

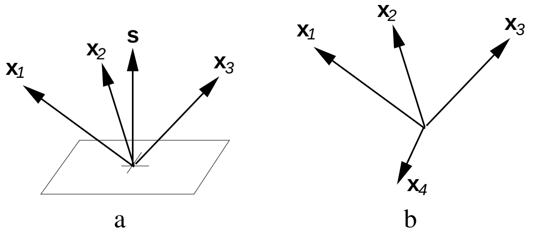


Figure 1. Directional (a) and panoramic (b) image.

A central linear camera is represented by its projection matrix $\mathbf{P} = (\mathbf{p}_1 \,|\, \mathbf{p}_2 \,|\, \mathbf{p}_3)^\top \in \mathbb{R}^{3\times4}, \operatorname{rank}\mathbf{P} = 3$. It projects a scene point $\mathbf{X}$ to the image point $\mathbf{x} = \mathbf{P}\mathbf{X}$. Two matrices $\mathbf{P}$ and $\mathbf{P}'$ represent the same camera iff $\exists \pi > 0 : \pi\mathbf{P} = \mathbf{P}'$. The *camera center* is defined as the mapping C that assigns the scene point $C(\mathbf{P}) \in \mathbb{R}^4$ to a camera $\mathbf{P}$ as [4, 3] $C_i(\mathbf{P}) = (-1)^i \, |\mathbf{P}_{[i]}|$ where C_i is the i-th component of C and $\mathbf{P}_{[i]} \in \mathbb{R}^{3\times3}$ is a matrix made from $\mathbf{P}$ by removing the i-th column, and $|.|$ denotes the determinant. In fact, $C(\mathbf{P})$ is the meet of the three planes represented by rows of $\mathbf{P}$, taking into account plane orientations. The 4-th component of the camera center $C_4(\mathbf{P}) = \Pi_\infty^\top C(\mathbf{P})$ is the determinant of the sub-matrix formed from the first three columns of $\mathbf{P}$. Its sign indicates the orientation of the camera coordinate system basis. $C_4(\mathbf{P}) > 0$ corresponds to a perspective camera yielding direct (meaning an opposite to mirrored) images, $C_4(\mathbf{P}) = 0$ corresponds to an orthographic camera, and $C_4(\mathbf{P}) < 0$ corresponds to a perspective camera yielding mirrored images.

Different scene and image points and cameras are distinguished by indices. Scene points are indexed by n in subscript and cameras are indexed by k in superscript. The geometry is assumed to consist of N scene points $\{\mathbf{X}_n \,|\, n = 1, \ldots, N\}$ and K cameras $\{\mathbf{P}^k \,|\, k = 1, \ldots, K\}$. Projections of all scene points by all cameras are assumed available, i.e., no image points are missing. The projection equations for all image points are $\mathbf{x}_n^k = \mathbf{P}^k\mathbf{X}_n$.

3. Weak and Strong Realizability

We will recall basics of Hartley's theory of weak and strong realizability [4, 3] and extend it for both panoramic and directional images.

It is known that a *projective reconstruction* from the set $\{\mathbf{x}_n^k\}$ of images exists iff the image points satisfy multi-view constraints [3]. However, not every projective reconstruction corresponds to any *real* geometry. A real geometry requires three additional constraints:

- *Positive scales.* Since we are in oriented projective space, scene points must be projected to image points using positive scales, $w^k_n \mathbf{x}^k_n = \mathbf{P}^k \mathbf{X}_n, w^k_n > 0$.

- *Scene points have positive depths.* As scene space is modeled by $A_3 = (T_3, \Pi_\infty)$, it must be $\mathbf{\Pi}_\infty^\top \mathbf{X}_n = X_{4n} > 0$.
- *The orientations of camera coordinate system bases in which image points are measured are positive.* It is assumed that the orientations of the true camera coordinate system bases are known. As negative orientations can be easily changed to positive, we assume without loss of generality all orientations positive, $\mathbf{\Pi}_\infty C(\mathbf{P}^k) = C_4(\mathbf{P}^k) > 0$

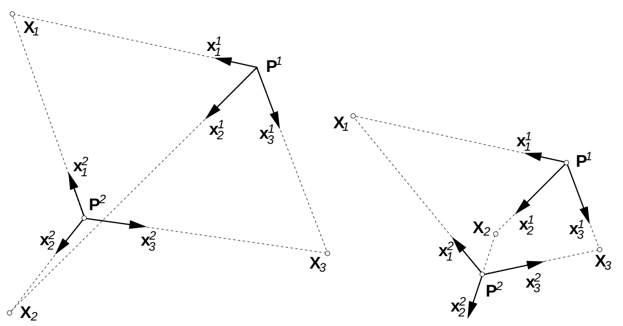
The fact that the class of images from which a projective reconstruction exists is larger than that from which a real geometry can be reconstructed is formalized by

Definition 2 *A weak realization of an image set* $\{\mathbf{x}^k_n\}$ *is the set* $\{\mathbf{P}^k, \mathbf{X}_n | \, \mathbf{X}_n \in \mathbb{R}^4, \mathbf{X}_n \neq \mathbf{0}, \, \mathbf{P}^k \in \mathbb{R}^{3\times 4}, \, \mathrm{rank}\, \mathbf{P}^k = 3\}$ *such that* $\forall k, n \, \exists w^k_n \neq 0 : \, w^k_n \mathbf{x}^k_n = \mathbf{P}^k \mathbf{X}_n$.

A strong realization of an image set $\{\mathbf{x}^k_n\}$ *is its weak realization* $\{\mathbf{P}^k, \mathbf{X}_n\}$ *such that* $\forall k, n : \, C_4(\mathbf{P}^k) > 0, X_{n4} > 0, \exists w^k_n > 0 : \, w^k_n \mathbf{x}^k_n = \mathbf{P}^k \mathbf{X}_n$.

An image set $\{\mathbf{x}^k_n\}$ *is called weakly resp. strongly realizable iff a weak resp. strong realization of it exists.*

A weak realization means what is usually called a projective reconstruction, and a strong realization means a weak realization with the above three constraints. In Fig. 2, (a) shows a strong realization $\{\mathbf{P}^k, \mathbf{X}_n\}$ of the image set $\{\mathbf{x}^k_n\}$ for $K = 2$ images and $N = 3$ scene points. In (b), the shown realization of the same image point set (illustrated by using the same vector triplets $\mathbf{x}^k_n$, the triplet in (b) is just moved and rotated in comparison with (a)) is weak because $w^2_2 < 0$. In both (a) and (b), it is assumed that $\mathbf{X}_{n4} > 0$ and $C_4(\mathbf{P}^k) > 0$, therefore only the signs of w^k_n determine whether the realization is weak or strong.



a b

Figure 2. Strong (a) and weak (b) realization.

If it exists, a strong realization can be computed from image correspondences $\{\mathbf{x}^k_n\}$ without knowing any camera calibration [4, 3]. It is done by a proper transformation of a weak realization $\{\mathbf{P}^k, \mathbf{X}_n\}$, computed by well-known algorithms for projective reconstruction [3]. Namely, $\{\mathbf{P}^k, \mathbf{X}_n\}$ is transformed to a strong realization $\{\bar{\mathbf{X}}_n, \bar{\mathbf{P}}^k | \, \xi_n \bar{\mathbf{X}}_n = \mathbf{H}\mathbf{X}_n \, , \pi^k \bar{\mathbf{P}}^k = \mathbf{P}^k \mathbf{H}^{-1}\}$, where the unknowns $\xi_n \neq 0, \pi^k \neq 0$ and the homography $\mathbf{H} \in \mathbb{R}^{4\times 4}$ are chosen as follows:

1. ξ_n and π^k have to satisfy the conditions $\xi_n \pi^k w^k_n > 0$ where $w^k_n \mathbf{x}^k_n = \mathbf{P}^k \mathbf{X}_n$.
2. $\mathbf{H}$ is chosen so that its 4-th row $\mathbf{h}_4$ and its determinant satisfy the system of inequalities $\mathbf{h}_4^\top \mathbf{X}_n > 0, \; |\mathbf{H}| \, \mathbf{h}_4^\top C(\mathbf{P}^k) > 0$ (*cheiral inequalities* [4, 3]).

The inequalities can be solved by linear programming. The solution is typically non-unique. Generally, there is no simpler way for testing the *existence* of a solution (and hence the strong realizability of an image set) than solving the inequalities. However, Hartley [4, 3] has proven the following necessary and sufficient condition for strong realizability of a pair of directional images. It enables to test for the strong realizability in a simpler way than by linear programming.

Theorem 1 *Let* $\{\mathbf{x}^k_n | \, k = 1, 2, \, x^k_{n3} > 0\}$ *be a pair of directional images and let* $\{\mathbf{P}^k, \mathbf{X}_n\}$ *be its arbitrary weak realization, i.e.,* $\forall k, n \, \exists w^k_n \neq 0 : w^k_n \mathbf{x}^k_n = \mathbf{P}^k \mathbf{X}_n$. *The image pair is strongly realizable if and only if* $w^1_n w^2_n$ *has equal signs for all* n.

The proof is in [4, 3] and its more compact version in [8]. The theorem generally holds neither for more than two images nor for two panoramic images.

Definition 2 is similar to that in [4, 3], however, it applies naturally for also panoramic rather than just directional images. The resulting cheiral inequalities are exactly the same. Hartley uses cheirality in the definition: a strong realization is defined as a weak realization such that $\forall k, n : \, \chi(\mathbf{P}^k, \mathbf{X}_n) > 0$. Yet cheirality $\chi(\mathbf{P}^k, \mathbf{X}_n)$ has no meaning for panoramic images, since the properties *in front of the camera* and *behind the camera* are not defined. A concept analogous to cheirality could be used but it would have to be defined for each image point separately, as e.g. $\chi(\mathbf{P}^k, \mathbf{X}_n, \mathbf{x}^k_n)$.

4. Conditions for Realizability of Two Images

Theorem 1 gives necessary and sufficient conditions for a strong realizability in terms of an existing weak realization. We state three theorems in this section, giving necessary and sufficient conditions for realizability in terms of epipoles, image points, fundamental matrix, and inter-image homography. The first one is the epipolar constraint $\mathbf{x}^\top \mathbf{F} \mathbf{x}' = 0$ with special handling of epipoles. The second one is the main result of this paper, the condition for a strong realizability of a pair of *directional* images. This simple condition is not sufficient for panoramic images. The condition for a strong realizability of a panoramic image pair is given by the third theorem. From now on, we will denote image points in the first and second image $\mathbf{x}, \mathbf{x}'$ respectively, rather than $\mathbf{x}^1, \mathbf{x}^2$. Similarly for camera matrices, epipoles, etc.

Weak realizability of an image pair $\{\mathbf{x}_n, \mathbf{x}'_n\}$ implies the epipolar constraint, yet the reversed implication is not true as the epipoles have to be treated in a special way. The following theorem captures this fact.

Theorem 2 *An image pair $\{\mathbf{x}_n, \mathbf{x}'_n\}$ is weakly realizable if and only if*

$$\begin{aligned}&\exists \mathbf{F} \in \mathbb{R}^{3\times 3}, \operatorname{rank}\mathbf{F} = 2, \forall n: \\ &\quad [(\mathbf{x}_n^\top \mathbf{F} \neq 0, \mathbf{F}\mathbf{x}'_n \neq 0) \vee (\mathbf{x}_n^\top \mathbf{F} = 0, \mathbf{F}\mathbf{x}'_n = 0)] \\ &\quad \wedge (\mathbf{x}_n^\top \mathbf{F}\mathbf{x}'_n = 0)\end{aligned} \tag{1}$$

Proof. The forward implication. Assume a weak realization of $\{\mathbf{x}_n, \mathbf{x}'_n\}$. It is well-known [3] that $\mathbf{P}' = (\mathbf{I}\,|\,\mathbf{0})$ can be assumed without loss of generality. Let

$$\begin{aligned}\mathbf{P} = (\mathbf{A}\,|\,\mathbf{e}), &\qquad \mathbf{P}' = (\mathbf{I}\,|\,\mathbf{0}) \\ \mathbf{X}_n = (\mathbf{x}'^\top_n\,|\,\rho_n)^\top, &\qquad w_n\mathbf{x}_n = \mathbf{A}\mathbf{x}'_n + \rho_n\mathbf{e}\end{aligned} \tag{2}$$

Vector-multiplying the fourth equation by $\mathbf{e}$ yields $w_n\mathbf{e} \times \mathbf{x}_n = \mathbf{e} \times \mathbf{A}\mathbf{x}'_n$ and, in turn, multiplying this by $\mathbf{x}_n^\top$ from the left results in $\mathbf{x}_n^\top \mathbf{e} \times \mathbf{A}\mathbf{x}'_n = 0$. Let $\mathbf{F} = \mathbf{e} \times \mathbf{A} = (\mathbf{e} \times \mathbf{a}_1\,|\,\mathbf{e} \times \mathbf{a}_2\,|\,\mathbf{e} \times \mathbf{a}_3)$ where $\mathbf{a}_i$ are columns of $\mathbf{A}$.

We show that $\operatorname{rank}\mathbf{F} = \operatorname{rank}(\mathbf{e} \times \mathbf{A}) = 2$. That says that the vectors $\mathbf{e} \times \mathbf{a}_i$ span a plane. They cannot span a 3-space because all must lie in a plane normal to $\mathbf{e}$. They cannot span a line because the vectors $\mathbf{a}_i$ span a plane or a 3-space, which is implied by the assumption $\operatorname{rank}(\mathbf{A}\,|\,\mathbf{e}) = 3$. They cannot be simultaneously zero since we can assume without loss of generality $\mathbf{e} \neq 0$, because if by chance $\forall n: w_n\mathbf{x}_n = \mathbf{A}\mathbf{x}'_n$, all $\rho_n = 0$ can be set to zero in (2).

The condition $(\mathbf{x}_n^\top \mathbf{F} \neq 0, \mathbf{F}\mathbf{x}'_n \neq 0) \vee (\mathbf{x}_n^\top \mathbf{F} = 0, \mathbf{F}\mathbf{x}'_n = 0)$ follows from $w_n\mathbf{e} \times \mathbf{x}_n = \mathbf{e} \times \mathbf{A}\mathbf{x}'_n$ because $w_n \neq 0$ and $(\mathbf{e} \times \mathbf{x} = 0) \Leftrightarrow (\mathbf{x}^\top \mathbf{F} = 0)$.

The backward implication. Assume that (1) holds. Any $\mathbf{F} \in \mathbb{R}^{3\times 3}$ of rank 2 can be decomposed as $\mathbf{F} = \mathbf{e} \times \mathbf{A}$ where $\mathbf{e} \in \mathbb{R}^3, \mathbf{e} \neq \mathbf{0}$ and $\mathbf{A} \in \mathbb{R}^{3\times 3}$, $\operatorname{rank}\mathbf{A} = 3$. We have $\mathbf{x}_n^\top \mathbf{e} \times \mathbf{A}\mathbf{x}'_n = 0$ meaning that the vectors $\mathbf{x}_n, \mathbf{A}\mathbf{x}'_n$ and $\mathbf{e}$ do not span a 3-space. The condition $\forall n: [(\mathbf{x}_n^\top \mathbf{F} \neq 0, \mathbf{F}\mathbf{x}'_n \neq 0) \vee (\mathbf{x}_n^\top \mathbf{F} = 0, \mathbf{F}\mathbf{x}'_n = 0)]$ ensures that the expressions $\mathbf{e} \times \mathbf{x}_n$ and $\mathbf{e} \times \mathbf{A}\mathbf{x}'_n$ are either both non-zero or both zero. Therefore it must be $\forall n\, \exists w_n, \rho_n \in \mathbb{R}: w_n\mathbf{x}_n = \mathbf{A}\mathbf{x}'_n + \rho_n\mathbf{e}$ and moreover $w_n \neq 0$. Letting $\mathbf{P} = (\mathbf{A}\,|\,\mathbf{e}), \mathbf{P}' = (\mathbf{I}\,|\,\mathbf{0})$ and $\mathbf{X}_n = (\mathbf{x}'^\top_n\,|\,\rho_n)^\top$ proves that the image pair $\{\mathbf{x}_n, \mathbf{x}'_n\}$ is weakly realizable. □

The following theorem gives a necessary and sufficient condition for a strong realizability of pair of directional images.

Theorem 3 *A pair $\{\mathbf{x}_n, \mathbf{x}'_n\,|\,x_{n3} > 0, x'_{n3} > 0\}$ of directional images is strongly realizable if and only if*

$$\begin{aligned}&\exists \mathbf{F} \in \mathbb{R}^{3\times 3}, \operatorname{rank}\mathbf{F} = 2, \exists \mathbf{e} \in \mathbb{R}^3, \mathbf{e} \neq \mathbf{0}, \mathbf{e}^\top\mathbf{F} = \mathbf{0}, \\ &\qquad \forall n\, \exists w_n > 0: \; w_n\mathbf{e} \times \mathbf{x}_n = \mathbf{F}\mathbf{x}'_n\end{aligned} \tag{3}$$

Proof. The forward implication. We use the intermediate result $w_n\mathbf{e} \times \mathbf{x}_n = \mathbf{e} \times \mathbf{A}\mathbf{x}'_n$ from the proof of Theorem 2. Since the realization is strong, it is $w_n > 0$.

The backward implication. Assume that (3) holds, i.e. $\forall n\, \exists w_n > 0: \; w_n\mathbf{e} \times \mathbf{x}_n = \mathbf{e} \times \mathbf{A}\mathbf{x}'_n$. Since the most general expression that vanishes under vector-multiplication by $\mathbf{e}$ is $\rho_n\mathbf{e}$ for any $\rho_n \in \mathbb{R}$, $\mathbf{e}$ can be separated to an extra term, i.e., $w_n\mathbf{x}_n = \mathbf{A}\mathbf{x}'_n + \rho_n\mathbf{e}$. Let $\mathbf{P} = (\mathbf{A}\,|\,\mathbf{e}), \mathbf{P}' = (\mathbf{I}\,|\,\mathbf{0})$ and $\mathbf{X}_n = (\mathbf{x}'^\top_n\,|\,\rho_n)^\top$. This is a strong realization since it satisfies the assumptions of Theorem 1. □

The simple condition given by Theorem 3 is sufficient only for directional images because Theorem 1 is used in the proof of the backward implication. The sufficient condition for general image pair is not so simple and given by

Theorem 4 *An image pair $\{\mathbf{x}_n, \mathbf{x}'_n\}$ is strongly realizable if and only if*

$$\begin{aligned}&\exists \mathbf{A} \in \mathbb{R}^{3\times 3}, |\mathbf{A}| > 0, \exists \mathbf{e} \in \mathbb{R}^3, \\ &\qquad \forall n\, \exists w_n > 0, \rho_n > 0: \; w_n\mathbf{x}_n = \mathbf{A}\mathbf{x}'_n + \rho_n\mathbf{e}\end{aligned} \tag{4}$$

Proof. The forward implication. Let $\{\mathbf{P}, \mathbf{P}', \mathbf{X}_n\}$ be a strong realization of $\{\mathbf{x}_n, \mathbf{x}'_n\}$. Let $\mathbf{P} = (\mathbf{M}\,|\,\mathbf{v}), \mathbf{P}' = (\mathbf{M}'\,|\,\mathbf{v}'), \mathbf{X}_n = (\mathbf{y}_n^\top\,|\,\rho_n)$ where $\mathbf{y}_n \in \mathbb{R}^3$ and $\rho_n \in \mathbb{R}$. Let $\mathbf{X}_n$ be so scaled that $\mathbf{P}'\mathbf{X}_n = \mathbf{x}'_n$. Then

$$w_n\mathbf{x}_n = \mathbf{M}\mathbf{y}_n + \rho_n\mathbf{v}, \quad \mathbf{x}'_n = \mathbf{M}'\mathbf{y}_n + \rho_n\mathbf{v}' \tag{5}$$

Expressing $\mathbf{y}_n$ from the second equation and substituting it to the first one yields $w_n\mathbf{x}_n = \mathbf{A}\mathbf{x}'_n + \rho_n\mathbf{e}$, where $\mathbf{A} = \mathbf{M}\mathbf{M}'^{-1}$ and $\mathbf{e} = -\mathbf{M}\mathbf{M}'^{-1}\mathbf{v}' - \mathbf{v}$. As $|\mathbf{M}| > 0, |\mathbf{M}'| > 0, \rho_n > 0$ and $w_n > 0$ due to the strong realizability, (4) is proven.

The backward implication. Assume (4) holds. Then $\mathbf{P} = (\mathbf{A}\,|\,\mathbf{e}), \mathbf{P}' = (\mathbf{I}\,|\,\mathbf{0})$ and $\mathbf{X}_n = (\mathbf{x}'^\top_n\,|\,\rho_n)^\top$ is the sought strong realization. □

5. Geometrical Insight

There is a *geometrical insight* behind the last three theorems. The conditions (1), (3) and (4) all say that the vectors $\mathbf{x}_n, \mathbf{A}\mathbf{x}'_n$ and $\mathbf{e}$ lay in a plane, expressing this fact respectively by the equations

$$\mathbf{x}_n^\top \mathbf{e} \times \mathbf{A}\mathbf{x}'_n = 0 \tag{6}$$

$$w_n\mathbf{e} \times \mathbf{x}_n = \mathbf{e} \times \mathbf{A}\mathbf{x}'_n = 0 \tag{7}$$

$$w_n\mathbf{x}_n = \mathbf{A}\mathbf{x}'_n + \rho_n\mathbf{e} \tag{8}$$

If just $w_n \neq 0$ is required, these three are equivalent. What makes the condition (3) stronger than (1), and the condition

(4) stronger than (3) are *signs of* w_n *and* ρ_n. Not considering the special handling of epipoles in Theorem 2 and the constraint $|\mathbf{A}| > 0$ in (4), the the necessary and sufficient conditions for weak realizability of two images (1), for strong realizability of two directional images (3), and for strong realizability of two images (4) be respectively expressed by the following three expressions:

$$\begin{aligned} \forall n\, \exists w_n \neq 0 \; : \; w_n \mathbf{x}_n &= \mathbf{A}\mathbf{x}'_n + \rho_n \mathbf{e} \quad (9) \\ \forall n\, \exists w_n > 0 \; : \; w_n \mathbf{x}_n &= \mathbf{A}\mathbf{x}'_n + \rho_n \mathbf{e} \quad (10) \\ \forall n\, \exists w_n > 0, \rho_n > 0 \; : \; w_n \mathbf{x}_n &= \mathbf{A}\mathbf{x}'_n + \rho_n \mathbf{e} \quad (11) \end{aligned}$$

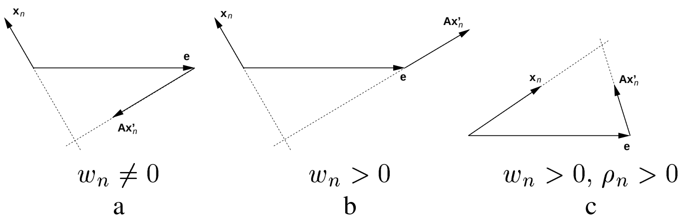


Figure 3. Geometrical meaning of the signs of w_n and ρ_n in the epipolar constraint $w_n\mathbf{x}_n = \mathbf{A}\mathbf{x}_n + \rho_n\mathbf{e}$.

Fig. 3 illustrates the geometrical meaning of (9), (10) and (11). The figure plane represent the epipolar plane corresponding to $\{\mathbf{x}_n, \mathbf{x}'_n\}$, in the coordinate system of the first camera. In each sub-figure, three vectors are depicted: the first image point $\mathbf{x}_n$, the first epipole $\mathbf{e}$, and the second image point transformed to the coordinate system of the first camera by the homography $\mathbf{A}$. They are free vectors, they can be placed arbitrarily. We placed them to make apparent what the numbers w_n, ρ_n must be so that $w_n\mathbf{x}_n = \mathbf{A}\mathbf{x}'_n + \rho_n\mathbf{e}$.

Fig. 3a shows an example of the ray triplet satisfying (9) and not satisfying (10) and (11). Nothing more is required than that the rays lay in a single plane. This is satisfied trivially in the figure because the figure plane itself is the actual epipolar plane. Fig. 3b shows an example of ray triplet satisfying (10) and not satisfying (11). The requirement $w_n > 0$ means that the vectors $\mathbf{x}_n$ and $\mathbf{A}\mathbf{x}'_n$ lay on the same side of $\mathbf{e}$, i.e. either both above $\mathbf{e}$ or both under $\mathbf{e}$. This is more conveniently apparent from the equivalent constraint $w_n\mathbf{e} \times \mathbf{x}_n = \mathbf{e} \times \mathbf{A}\mathbf{x}'_n, w_n > 0$. Fig. 3c shows an example of ray triplet satisfying (9). The requirement $w_n > 0, \rho_n > 0$ means that the half-lines $\mathbf{x}_n$ and $\mathbf{A}\mathbf{x}'_n$ *intersect*. This corresponds to the requirement that the camera rays intersect in the scene.

Note how the knowledge about ray orientations is lost during the derivation of the epipolar constraint. Let us consider a strongly realizable image pair. According to Theorem 4, the strong realizability is equivalent to (4). Epipolar constraint $\mathbf{x}_n^\top \mathbf{F}\mathbf{x}'_n = 0$ is derived from (4) in the following two steps: (i) the equation $w_n\mathbf{x}_n = \mathbf{A}\mathbf{x}'_n + \rho_n\mathbf{e}$ is vector-multiplied from the left by $\mathbf{e}$, canceling the term $\rho_n\mathbf{e}$, and (ii) the resulting equation $w_n\mathbf{e} \times \mathbf{x}_n = \mathbf{e} \times \mathbf{A}\mathbf{x}'_n$ is multiplied by $\mathbf{x}_n^\top$ from the left, canceling the term $w_n\mathbf{e} \times \mathbf{x}_n$. The signs of ρ_n are lost after the first step and the signs of w_n are lost after the second step. The knowledge before the first step, after the first step, and after the second step is expressed respectively by (11), (10) and (9).

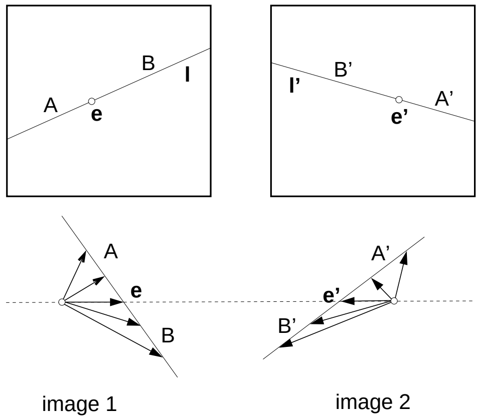


Figure 4. The condition sufficient for directional images but insufficient for panoramic images.

Consider two directional images in Fig. 4 with epipoles $\mathbf{e}, \mathbf{e}'$ and a pair $\mathbf{l}, \mathbf{l}'$ of corresponding epipolar lines. A, B and A', B' denote sets of image points on $\mathbf{l}$ and $\mathbf{l}'$ respectively, where all points in one set lay on a single side of the epipole. Clearly, the *necessary* condition for strong realizability is that either A correspond to A' and B to B', or A corresponds to B' and B to A'. This is what (10) expresses. The ambiguity is caused by not knowing whether an epipole is in front or behind the image plane [5].

Why the condition is sufficient for directional and insufficient for panoramic images? For directional images, once (10) is satisfied and (11) not, we can always find some inter-image homography $\bar{\mathbf{A}}$, different from $\mathbf{A}$ but also compatible with the given $\mathbf{F}$, such that (11) is satisfied, i.e., the rays intersect. In Fig. 4, the changing of the inter-image homography can be imagined as slanting the retinas with fixed image points in such a way that the epipoles and camera centers remain collinear. However, this is not possible for panoramic cameras in which no such planar retina exists.

6. Applications

Theorem 3 is a main result of the paper. We outline two applications in this section.

Assume finding correspondences in two directional images by a stereo algorithm. If the plane at infinity is known, the search space can be reduced by the requirement that all scene points are 'not further away than infinity' [3]. This is not possible for weakly calibrated cameras (because the plane at infinity is unknown), however, the search space can *sometimes* be reduced by applying Theorem 3. More exactly, it can be reduced only if the epipole *separates* the set of correspondence candidates on at least one epipolar line pair in at least one image. In other words, if the candidates are in at least three of the sets A, B, A', B' in Fig. 4. If the candidates are e.g. only in A and A', (3) is satisfied for all their configurations, i.e., all their configurations are strongly realizable. This occurs if the epipoles are outside the images, which is frequent when doing stereo.

For two weakly calibrated panoramic cameras, Theorem 4 would have to be used to reduce the search space. Using it is much less straightforward, as the inter-image homography is present in (4). The reason is that a given fundamental matrix obtained by weak calibration cannot be uniquely decomposed to $\mathbf{e} \times \mathbf{A}$, as it is $\mathbf{e} \times \mathbf{A} = \mathbf{e} \times (\mathbf{A} + \mathbf{e}\omega^\top)$ for any $\omega \in \mathbb{R}^3$. In fact, testing the strong realizability would lead to linear programming, similarly as solving cheiral inequalities.

The second application is ruling out impossible solutions for fundamental matrix. When a fundamental matrix is computed from seven points, one or three solutions are obtained, corresponding to the roots of the cubic equation enforcing the constraint $\operatorname{rank} \mathbf{F} = 2$ [3]. For example, when using RANSAC technique for robust estimation of epipolar geometry [2], Theorem 4 allows to rule out one or two of the three solutions in advance in each sample.

Unfortunately, (4) does not ensure that only the correct single solution for $\mathbf{F}$ is always left. We observed in numerical simulations that one, two or three of the original three solutions are left. Approx. $30 - 50\%$ of the incorrect solutions were recognized using (4) in a realistic situation.

7. Conclusion

We extended Hartley's cheirality theory for panoramic central linear cameras. These are cameras possessing a single projection center, such that the transformation between scene and image points is linear in homogeneous coordinates, and having the field of view 360° at least in one great circle. The conventional cameras are called directional. For a directional camera, there exist a plane such that all image point vectors are on its single side. No such plane exists for a panoramic camera.

We focused on the case of two cameras. For both panoramic cameras, the well-known epipolar constraint holds. However, not all image points satisfying $\mathbf{x}^\top \mathbf{F} \mathbf{x}' = 0$ correspond to any real geometry. The reason is that the cheirality constraints must be considered too. We have presented necessary and sufficient conditions for an image point set to correspond to any real geometry. These conditions are formulated in terms of image points, epipoles, fundamental matrix, and inter-image homography.

There are different conditions for directional and panoramic cameras. The directionality brings an important additional knowledge that makes the sufficient conditions simpler and much more straightforward to use. For directional images, this condition is that either all epipolar lines aim towards the epipole or all aim out of it in each image.

We have presented the insight to the epipolar geometry, based on these theorems. In short, the conditions differ in knowledge of the signs of w_n and ρ_n in the epipolar equation expressed in terms of inter-image homography and projective depth, $w_n \mathbf{x}_n = \mathbf{A}\mathbf{x}'_n + \rho_n \mathbf{e}$.

Possible applications are the reduction of the search space and ruling out impossible matches in stereo, and ruling out impossible solutions for a fundamental matrix computed from seven points.

References

[1] Simon Baker and Shree K. Nayar. A theory of single-viewpoint catadioptric image formation. *International Journal of Computer Vision*, 35(2):1 – 22, 1999.

[2] Paul Beardsley, Phil Torr, and Andrew Zisserman. 3D model acquisition from extended image sequences. In Bernard Buxton and Roberto Cippola, editors, *European Conference on Computer Vision 1996*. Springer-Verlag, 1996. in version of technical report.

[3] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, UK, 2000.

[4] Richard I. Hartley. Chirality. *Int. Jour. Computer Vision IJCV*, 26(1):41–61, 1998.

[5] S. Laveau and O. Faugeras. Oriented projective geometry for computer vision. In *Proceedings of the 4th European Conference on Computer Vision'96*, volume I, pages 147–156. Springer-Verlag, April 1996.

[6] Tomáš Pajdla, Tomáš Svoboda, and Václav Hlaváč. Epipolar geometry of central panoramic cameras. In Ryad Benosman and Sing Bing Kang, editors, *Panoramic Vision : Sensors, Theory, and Applications*. Springer Verlag, Berlin, Germany, 1 edition, 2000.

[7] Jorge Stolfi. *Oriented Projective Geometry: A Framework for Geometric Computations*. Academic Press, Inc., 1250 Sixth Avenue, San Diego, CA 92101, 1991.

[8] Tomáš Werner and Tomáš Pajdla. Cheirality in epipolar geometry. Research Report CAK-340-03-1-2000-01, CTU-CMP-2000-21, Center for Machine Perception, Czech Technical University, Prague, Czech Republic, December 2000.

[9] Tomáš Werner, Tomáš Pajdla, and Václav Hlaváč. Oriented projective reconstruction. In M. Gengler, M. Prinz, and E. Schuster, editors, *Pattern Recognition and Medical Computer Vision: 22-nd Workshop of the Austrian Association for Pattern Recognition (ÖAGM/IAPR)*, pages 245–254, Wien, Austria, May 14–15 1998. Österreichische Computer Gesselschaft.